DOCUMENT RESUME

ED 093 960 TM 003 778

AUTHOR Halderson, Judith S.; Glasnapp, Douglas R.
TITLE Error Rates of Multiple F Tests in Factorial ANOVA Designs.
PUB DATE Apr 74
NOTE 24p.; Paper presented at American Educational Research Association Annual Meeting (Chicago, Illinois, April, 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS *Analysis of Variance; Computer Programs; Error Patterns; *Hypothesis Testing; Sampling; *Tests of Significance

ABSTRACT

The primary purpose of the present study was to investigate empirically the effect of multiple hypothesis testing on error rates in factorial ANOVA designs under a variety of controlled conditions. The per comparison, per experiment, and experimentwise error rates were investigated for three hypothesis testing procedures. The specific conditions manipulated included: (1) the number of factors in the design, (2) the number of levels of each factor, (3) the number of observations per cell, and (4) the population values of each null hypothesis (magnitude of the effect). A Monte Carlo computer simulation procedure was used for generation of the data. Type I and Type II errors were tabulated where appropriate fof the three error rates. (Author)

ERROR RATES OF MULTIPLE *F* TESTS IN FACTORIAL ANOVA DESIGNS

Judith S. Halderson
National Assessment of Educational Progress

and

Douglas R. Glasnapp
University of Kansas

Paper Presented at the Annual Meeting of the American Educational Research Association, Chicago, 1974.

Error Rates of Multiple *F* Tests In Factorial ANOVA Designs

The concept of error rates in hypothesis testing was introduced to help the experimenter estimate the frequency of erroneous inferences. With one statistical test, error rates can be estimated if one knows the significance level and direction of the test as well as the sample size used. With more than one statistical test, two other factors must be considered: the number of hypotheses tested and the dependence among the tests.

The number of hypotheses tested in an experiment is extremely important in multiple hypothesis testing because of the accumulation of errors. For example, with 100 independent tests of true null hypotheses, each tested at the .05 level of significance, it would be expected that 5 of the tests would be significant just by chance alone. Error rates are also affected by the dependence among the tests, such as in tests of all pair-wise comparisons based on a set of independent means.

The problems of multiple hypothesis testing and their effects on error rates have been thoroughly investigated for the case of several independent means such as in the context of one-way analysis of variance (ANOVA) designs (Tukey, 1953; Ryan, 1959; Games, 1971). The same problems of multiple hypothesis testing also exist with higher-order ANOVA designs.

In complex ANOVA designs, many hypotheses may be may be tested within a single experiment. For example, in a five-factor fixed-effects design, there may be as many as 31 hypotheses. With each hypothesis tested at the .05 level of significance, and assuming all null hypotheses are true, it would be expected that an experimenter would be making 1.5 Type I errors per experiment. Higher-order ANOVA designs also have the problem of dependence among the tests. Although it is true that the numerator sums of squares are based on independent components of the total sum of squares (Lindquist, 1953), the tests themselves may be based on the same mean square for error; thus, there may be dependence among the tests which could affect the frequency of errors within the total design (Hays, 1963).

At the present time, very little is known about the effects of multiple hypothesis testing on error rates in multi-dimensional ANOVA designs. Empirical research has centered on the problem of comparisons based on a set of independent means as in one-way designs (Petrinovich & Hardyck, 1969; Norton & Bulgren, 1965). Theoretical discussions of an appropriate error rate to describe the frequency of errors in multiple testing situations have also centered on the one-way context (Tukey, 1953; Ryan, 1959; Games, 1971).

Purpose

The purpose of this study was to empirically investigate the frequency of erroneous conclusions in factorial ANOVA designs under a variety of controlled conditions.

The frequency of errors was measured using the following three error rates: the error rate per test, the error rate per experiment, and the experimentwise error rate (Tukey, 1953; Hartley, 1955; Ryan, 1959). Since the method of testing each hypothesis affects the frequency of erroneous conclusions, three different hypothesis testing procedures were used: (a) test each hypothesis at a specified alpha (α) level of significance, (b) use Hartley's (1955) sequential testing procedure[1] designed to control the experimentwise error rate at a specified α level, and (c) test each hypothesis at the $\alpha/\underline{k}$ level for $\underline{k}$ tests according to a Bonferroni procedure (Miller, 1966; Games, 1971). Factorial designs were varied according to the number of factors, the number of levels of each factor, the number of observations per cell, and the population values of the null hypotheses (all true, all false, and combinations of both true and false in the same design). Where appropriate, Type I and/or Type II error rates were calculated. In all cases, α was set at .05.

Procedure

Both two- and three-way completely crossed fixed effects factorial designs with independent groups per cell were studied. The designs selected were the 2x2, 2x3, 2x4, 2x5, 5x5, 2x2x2, and 5x5x5 designs. For all designs, the number of observations per cell were equal.

The data were randomly generated by computer using a Monte Carlo simulation procedure.[2] All data were drawn from normal distributions with equal variances and were generated according to the general linear model for ANOVA designs. The values of the main and interaction effects were calculated using Cohen's (1969) _f_ index of effect size. The magnitudes of effects were varied across four points: zero (_f_ = .00), small (_f_ = .10), medium (_f_ = .25), and large (_f_ = .40). For all designs, the main and interaction effect sizes were held constant. In addition, for the 2x4 and 2x2x2 designs effect size was varied across main and interaction effects.

A single simulation procedure consisted of the generation of one set of scores for a single design under combinations of the following conditions: (a) dimensions of the design, (b) cell size, and (c) population value of each effect. Acceptance or rejection of each hypothesis was determined using three criteria: (a) the usual _F_ procedure with α = .05 as the significance level for each test, (b) Hartley's sequential procedure, and (c) a Bonferroni procedure with .05/_k_ as the significance level for each test. Each simulation procedure was repeated 2000 times. Following the 2000 replications, the following three error rates were calculated for each of the hypothesis testing procedures: (a) per comparison error rate or the average of the individual hypothesis error rates, (b) per experiment error rate or the average number of errors per experiment, and (c) the experimentwise error rate or the proportion of experiments with at least one error.

Results

Per comparison error rates. Table 1 presents the per comparison error rates for the true null condition (all f's = .00 in a given design). The per comparison error rates of the alpha procedure fluctuated around the nominal .05 level with a median value of .0520. When α/k was used to test each hypothesis, the per comparison error rates were close to the expected values: the median per comparison error rate for two-way designs was .0172 (nominal level of .05/3 = .0167); the median was .0071 for three-way designs (nominal level of .05/7 = .0071). The per comparison error rates obtained by Hartley's procedure were just slightly higher than those obtained using the Bonferroni procedure. For two-way designs, the median error rate using Hartley's procedure was .0178; with three-way designs, the median was .0072.

The per comparison error rates for all null hypotheses false (all f's $>$ .00) are given in Table 2. Under these conditions, the per comparison error rate is the average of the Type II error rates of the individual hypotheses. One minus the per comparison error rate is the average power of the tests.

With all three hypothesis testing procedures, the Type II error rates were affected by the total sample size and the size of the effects in the design. As the sample size increased, either by increasing the cell size or by increasing the number of levels or factors with a constant cell size, the per comparison error rates declined. As the magnitude of the effects increased, the number of errors decreased. That is, all three procedures were the most powerful with large sample sizes and large effects present. With small sample sizes and small effects, all three procedures had almost equal difficulty detecting small deviations from the true null condition. In the most extreme case of large sample sizes and large effects (5x5 design with $\underline{n}$ = 30 and all effects defined by $\underline{f}$ = .40), all procedures were equally capable of rejecting the null hypotheses.

In general, the fewest Type II errors were made using the alpha procedure; the most were made using $\alpha/\underline{k}$. The frequency of errors made using Hartley's procedure was closer to that obtained using $\alpha/\underline{k}$ when effect sizes were small but approached the Type II error rate of the alpha procedure as effect size and sample sizes increased.

When both true and false null hypothese were present in the same design, both Type I and Type II per comparison error rates were calculated. The results are presented in Table 3 for the 2x4 designs and Table 4 for the 2x2x2 designs. The per comparison error rates were calculated as averages across all hypotheses. The frequency of Type I errors closely followed the previous results: the alpha procedure returned the most Type I errors and the $\alpha/\underline{k}$ procedure returned the fewest. Because of the sequential nature of Hartley's procedure, the status of all hypotheses affected the average error rates. Thus, as the proportion of false null hypotheses increased, the Type I error rate of Hartley's procedure approached the level obtained by the alpha procedure.

Per experiment error rates. The per experiment error rate is the average number of errors per experiment. It is also equal to the number of hypotheses times the per comparison error rate. Because of this direct relationship with the per comparison error rates, the per experiment error rates will not be discussed in detail nor reproduced here.[3]

For all true null hypotheses, the average number of errors per experiment was close to *k* times the nominal level of significance used to test each hypothesis. For the two-way designs, the per experiment rates for the $\alpha = .05$ procedure were close to .15. Using .05/3, the error rates were close to .05. For three-way designs, the average number of errors using $\alpha = .05$ was close to .35. The Bonferroni procedure maintained a per experiment rate of close to .05. For all designs, Hartley's procedure produced just slightly more Type I errors than the Bonferroni procedure.

The results obtained when all effects were non-zero and when the effects varied in a design paralleled those presented on per comparison error rates.

Experimentwise error rates. The experimentwise error rates when all effects were set at zero are presented in Table 5. These error rates are equal to the proportion of experiments with at least one Type I error.

When all null hypothese were true, the experimentwise error rates using Hartley's procedure and the Bonferroni procedure were identical. That is, if one hypothesis were rejected using the $\alpha/\underline{k}$ level in the Bonferroni procedure, then that same hypothesis was also rejected by Hartley's procedure, which begins its sequential testing at the $\alpha/\underline{k}$ level. Both of these procedures controlled the experimentwise error rates at the .05 level for both two- and three-way designs. When $\alpha = .05$ was used to test each hypothesis, the frequency of errors increased as the number of tests increased. With three hypotheses in the two-way designs, the experimentwise error rates had a median value of .1480. With seven tests in three-way designs, the median experimentwise error rate of the alpha procedure jumped to .2970.

With all null hypotheses false, the experimentwise error rate was equal to the proportion of experiments with at least one Type II error (see Table 6). With Type II errors, all testing procedures had different experimentwise error rates. The Type II error rates followed the same trends as the Type II per comparison error rates. In all cases, the alpha procedure had the lowest experimentwise error rate and the Bonferroni procedure had the highest rate. The power of all three procedures increased as the total number of observations and the sizes of the effects increased.

With both true and false null hypotheses in the same design, two experimentwise error rates were calculated: the proportion of experiments with at least one Type I error and the proportion with at least one Type II error. These error rates are presented in Tables 7 and 8. All error rates were affected by the number of false null hypotheses and followed trends already noted for Type I and II experimentwise error rates. As with the per comparison error rates, the experimentwise error rates of Hartley's procedure were complicated by the sequential nature of the test.

Discussion

This study has empirically investigated the frequency of Type I and Type II errors in selected factorial ANOVA designs. The frequency of errors was measured by three different error rates: the per comparison error rates or the average of the individual hypothesis rates and two experiment-based error rates, the per experiment error rates and the experimentwise error rates.

When the significance level is set at .05 for each hypothesis, a practice commonly followed in educational research, the accumulation of Type I errors as measured by the per experiment and experimentwise error rates was readily apparent. In this study, the per experiment error rates for all designs were close to $\underline{k}$ times .05 for $\underline{k}$ tests. The experimentwise error rates were close to .15 and .30 for two- and three-way designs, respectively. It should be noted that these experimentwise error rates are close to the expected experimentwise error rates of $\underline{k}$ independent tests as estimated by the formula $1 - (1 - \alpha)^{\underline{k}}$ where α is the level of significance used for each hypothesis.

Thus, it appears that the dependence among the tests in factorial designs when the same mean square for error is used with each test has little, if any, effect on the experimentwise error rates. In addition, the estimate based on *k* independent tests appears to be valid with small as well as large degrees of freedom for error. The alpha procedure, then, controls the individual hypothesis error rates but allows the experiment-based error rates to increase as the number of tests increases.

The Bonferroni and Hartley procedures adjusted the individual hypothesis significance levels to maintain the experiment-based error rates at acceptable and pre-specified levels. The Bonferroni procedure used α/k for the significance level of each hypothesis while Hartley's procedure used different significance levels for each hypothesis based on their obtained *p* values. The individual hypothesis rates of both procedures were close to α/k, thus adjusting the individual hypothesis rates as *k* increased. The per experiment and experimentwise error rates of these two procedures were held at the .05 level in both two- and three-way designs.

The frequency of Type II errors depended on the total number of observations and the sizes of the effects present. All three procedures had difficulty detecting small effects with small sample sizes; all were capable of detecting large effects with large sample sizes.

The choice of an appropriate hypothesis testing technique depends on the importance of experiment-based errors. If individual hypotheses are of major concern rather than a pattern of results obtained from an entire design, then a procedure which controls the individual hypothesis error rates at acceptable levels, such as the alpha procedure used in the present study, would be appropriate. However, if the accumulation of possible errors across an entire design would severely limit the validity of an experiment, then a procedure that controls the experiment-based error rates, such as a Bonferroni procedure or Hartley's procedure for factorial designs, would be more appropriate. The choice between individual hypothesis error rates and experiment based error rates has been debated elsewhere (Miller, 1966; Ryan, 1962; Wilson, 1962; Petrinovich & Hardyck, 1969) with no apparent resolution. Regardless of which type of error an experimenter chooses to control, he should be aware of the accumulation of errors that result from multiple hypothesis testing.

References

Cohen, J. *Statistical power analysis for the behavioral sciences.* New York: Academic Press, 1969.

Games, P. A. Multiple comparisons of means. *American Educational Research Journal*, 1971, *8*, 531-565.

Halderson, J. S. An empirical investigation of error rates and measures of association in factorial analysis of variance designs. Unpublished doctoral dissertation, University of Kansas, 1973.

Hartley, H. O. Some recent developments in analysis of variance. *Communications on Pure and Applied Mathematics*, 1955, *8*, 47-72.

Hays, W. L. *Statistics*. New York: Holt, Rinehart & Winston, 1963.

Lindquist, E. F. *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin, 1953.

Miller, R. G. *Simultaneous statistical inference*. New York: McGraw-Hill, 1966.

Norton, D. W., & Bulgren, W. G. A sampling study of certain multiple comparison procedures. Paper presented at the meeting of the American Educational Research Association, Chicago, February, 1965.

Petrinovich, L. F., & Hardyck, C. D. Error rates for multiple comparison methods: Some evidence concerning the frequency of erroneous conclusions. *Psychological Bulletin*, 1969, *71*, 43-54.

Ryan, T. A. Multiple comparisons in psychological research. *Psychological Bulletin*, 1959, *56*, 26-47.

Ryan, T. A. The experiment as the unit for computing rates of error. *Psychological Bulletin*, 1962, *59*, 301-305.

Tukey, J. W. The problem of multiple comparisons. Unpublished manuscript. Princeton University, 1953.

Wilson, W. A note on the inconsistency inherent in the necessity to perform multiple comparisons. *Psychological Bulletin*, 1962, *59*, 296-300.

Footnotes

1. Hartley's (1955) procedure tests each hypothesis in a factorial design based on its obtained *p* value. All hypotheses are ranked according to their *p* values and the hypothesis with the smallest *p* value is tested for significance at the α/k level where *k* is the number of hypotheses in the design. If the hypothesis is significant, the next ranking hypothesis is tested at the $\alpha/(k-1)$ level. The testing continues until a hypothesis is not rejected. At that point, testing stops and all remaining hypotheses are declared non-significant. Hartley demonstrates how this sequential procedure maintains an experimentwise error rate at α.

2. A more detailed description of the computer simulation procedure, including information concerning the random number generator used, can be found in Halderson, 1973.

3. Tables of per experiment rates can be found in Halderson, 1973.

Table 1

Per Comparison Error Rates

All f's = 0

Design	n	Procedure		
		Alpha[a]	Bonferroni[b]	Hartley[c]
2x2	5 15 30	.0572 .0583 .0498	.0183 .0180 .0187	.0195 .0188 .0190
2x3	5 15 30	.0455 .0542 .0528	.0152 .0187 .0178	.0153 .0193 .0178
2x4	5 15 30	.0543 .0520 .0523	.0190 .0155 .0162	.0200 .0157 .0163
2x5	5 15 30	.0502 .0490 .0502	.0180 .0155 .0155	.0188 .0162 .0158
5x5	5 15 30	.0457 .0533 .0547	.0145 .0172 .0172	.0150 .0180 .0177
2x2x2	5 15 30	.0515 .0488 .0530	.0077 .0054 .0071	.0079 .0054 .0071
5x5x5	5	.0514	.0071	.0072

[a]Nominal level of .05 for each hypothesis.

[b]Nominal level of .05/k for each hypothesis; k is the number of tests in any one experiment.

[c]Nominal level for each hypothesis is not known.

Table 2

Per Comparison Error Rates

All f's > 0

Design	f	n	Error Rates		
			Alpha[a]	Bonferroni[b]	Hartley[c]
2x2	.10	5	.9347	.9747	.9738
		15	.8863	.9557	.9537
		30	.8265	.9160	.9125
	.25	5	.8338	.9275	.9212
		15	.5788	.7395	.7067
		30	.3060	.4685	.3818
	.40	5	.6575	.8142	.7880
		15	.2255	.3672	.2785
		30	.0492	.0952	.0508
2x3	.10	5	.9245	.9712	.9698
		15	.8782	.9455	.9430
		30	.7975	.8992	.8923
	.25	5	.8065	.9073	.8988
		15	.4912	.6557	.6113
		30	.2155	.3425	.2553
	.40	5	.5910	.7502	.7210
		15	.1505	.2498	.1702
		30	.0238	.0505	.0248
2x4	.10	5	.9272	.9742	.9723
		15	.8700	.9438	.9417
		30	.7608	.8692	.8600
	.25	5	.7755	.8895	.3767
		15	.4215	.5795	.5185
		30	.1712	.2670	.1945
	.40	5	.5120	.6753	.6282
		15	.1075	.1855	.1147
		30	.0098	.0263	.0098

[a]Nominal level of .05 for each hypothesis.

[b]Nominal level of .05/k for each hypothesis; k is the number of tests in any one experiment.

[c]Nominal level for each hypothesis is not known.

Table 2

(Continued)

Design	$\underline{f}$	$\underline{n}$	Error Rates		
			Alpha	Bonferroni	Hartley
2x5	.10	5	.9200	.9717	.9700
		15	.8448	.9272	.9247
		30	.7313	.8562	.8435
	.25	5	.7402	.8632	.8503
		15	.3738	.5277	.4560
		30	.1302	.2112	.1415
	.40	5	.4537	.6068	.5535
		15	.0693	.1318	.0717
		30	.0055	.0145	.0055
5x5	.10	5	.8937	.9578	.9565
		15	.7533	.8618	.8505
		30	.5507	.6930	.6573
	.25	5	.5435	.7002	.6615
		15	.1207	.1930	.1297
		30	.0087	.0233	.0087
	.40	5	.1823	.2778	.2050
		15	.0022	.0075	.0022
		30	.0000	.0000	.0000
2x2x2	.10	5	.9154	.9854	.9852
		15	.8491	.9634	.9623
		30	.7443	.9184	.9136
	.25	5	.7458	.9246	.9198
		15	.4071	.6840	.6286
		30	.1723	.3746	.2601
	.40	5	.4893	.7569	.7209
		15	.1143	.2767	.1618
		30	.0236	.0731	.0251

Table 3

Per Comparison Error Rates:

2x4 Designs with Various *f*'s

				Type I			Type II		
f_A	f_B	f_{AB}	*n*	Alpha[a]	Bonf.[b]	Hartley[c]	Alpha[a]	Bonf.[b]	Hartley[c]
.25	0	0	5	.0340	.0115	.0130	.2207	.2710	.2707
			15	.0310	.0092	.0133	.0763	.1272	.1272
			30	.0323	.0128	.0173	.0097	.0233	.0232
0	.25	0	5	.0358	.0117	.0130	.2620	.3002	.2993
			15	.0320	.0102	.0120	.1315	.1955	.1950
			30	.0355	.0117	.0168	.0267	.0567	.0560
.25	0	.25	5	.0170	.0057	.0070	.5087	.5823	.5793
			15	.0178	.0073	.0105	.2962	.3930	.3795
			30	.0160	.0047	.0122	.1432	.2073	.1883
.40	0	.25	5	.0172	.0052	.0088	.3862	.4712	.4637
			15	.0167	.0065	.0098	.2275	.2810	.2665
			30	.0187	.0048	.0147	.1303	.1847	.1653
0	.25	.25	5	.0165	.0057	.0062	.5518	.6168	.6143
			15	.0137	.0052	.0067	.3557	.4630	.4498
			30	.0192	.0067	.0130	.1595	.2452	.2225
0	.40	.25	5	.0160	.0052	.0072	.4562	.5462	.5420
			15	.0192	.0065	.0112	.2320	.2903	.3738
			30	.0180	.0063	.0143	.1265	.1865	.1643

[a]Nominal level of .05 for each hypothesis.

[b]Nominal level of .05/*k* for each hypothesis; *k* is the number of tests in any one experiment.

[c]Nominal level for each hypothesis is not known.

Table 4

Per Comparison Error Rates:

2x2x2 Designs with Various $\underline{f}$'s

				Type I			Type II		
$\underline{f}_A$	$\underline{f}_{AB}$	$\underline{f}_{ABC}$[a]	$\underline{n}$	Alpha[b]	Bonf.[c]	Hartley[d]	Alpha[b]	Bonf.[c]	Hartley[d]
.25	0	0	5	.0441	.0053	.0054	.0930	.1254	.1254
			15	.0465	.0060	.0068	.0316	.0719	.0717
			30	.0475	.0054	.0062	.0049	.0181	.0181
.25	.25	0	5	.0366	.0049	.0051	.2106	.2626	.2622
			15	.0414	.0054	.0063	.1042	.1825	.1799
			30	.0404	.0056	.0076	.0351	.0899	.0854
.40	.25	0	5	.0389	.0054	.0057	.1551	.2202	.2194
			15	.0364	.0056	.0074	.0739	.1193	.1171
			30	.0381	.0049	.0074	.0292	.0699	.0656
.25	.25	.25	5	.0301	.0032	.0034	.3321	.4001	.3991
			15	.0272	.0039	.0046	.2078	.3119	.3078
			30	.0344	.0040	.0050	.1051	.1979	.1889
.40	.40	.25	5	.0304	.0040	.0052	.3517	.3482	.3459
			15	.0279	.0039	.0056	.1248	.1899	.1823
			30	.0310	.0045	.0078	.0741	.1190	.1129

[a]All additional effects set at zero; that is,

$\underline{f}_B = \underline{f}_C = \underline{f}_{AC} = \underline{f}_{BC} = 0.$

[b]Nominal level of .05 for each hypothesis.

[c]Nominal level of .05/$\underline{k}$ for each hypothesis; $\underline{k}$ is the number of tests in any one experiment.

[d]Nominal level for each hypothesis is not known.

Table 5

Experimentwise Error Rates

All $\underline{f}$'s = 0

Design	$\underline{n}$	Procedure: Alpha[a]	Procedure: Bonferroni/ Hartley[b]
2x2	5	.1550	.0535
	15	.1640	.0535
	30	.1425	.0550
2x3	5	.1275	.0440
	15	.1545	.0545
	30	.1525	.0535
2x4	5	.1480	.0555
	15	.1470	.0455
	30	.1490	.0480
2x5	5	.1445	.0530
	15	.1385	.0460
	30	.1420	.0460
5x5	5	.1255	.0435
	15	.1515	.0515
	30	.1545	.0510
2x2x2	5	.2920	.0500
	15	.2955	.0370
	30	.3160	.0490
5x5x5	5	.2985	.0485

[a]Nominal level of .05 for each hypothesis; experimentwise rate not known.

[b]Experimentwise error rates are identical and equal to .05 using these two methods when all $\underline{f}$'s = 0. Bonferroni uses a nominal level of $.05/\underline{k}$ for each hypothesis. Nominal level for each hypothesis for Hartley is unknown.

Table 6

Experimentwise Error Rates

All *f*'s > 0

Design	*f*	*n*	Error Rates		
			Alpha[a]	Bonferroni[b]	Hartley[c]
2x2	.10	5	.9995	.9995	.9995
		15	.9955	.9990	.9985
		30	.9960	.9995	.9980
	.25	5	.9905	.9985	.9925
		15	.9280	.9795	.9400
		30	.6885	.8680	.7075
	.40	5	.9475	.9845	.9555
		15	.5480	.7590	.5675
		30	.1465	.2755	.1475
2x3	.10	5	.9975	.9995	.9980
		15	.9975	.9995	.9975
		30	.9965	1.0000	.9970
	.25	5	.9875	.9980	.9900
		15	.8815	.9600	.8985
		30	.5485	.7660	.5655
	.40	5	.9250	.9745	.9365
		15	.4110	.6125	.4190
		30	.0715	.1485	.0725
2x4	.10	5	.9990	1.0000	.9995
		15	.9975	.9995	.9990
		30	.9910	.9990	.9940
	.25	5	.9845	.9975	.9880
		15	.8285	.9425	.8440
		30	.4735	.6660	.4835
	.40	5	.8885	.9610	.8975
		15	.3080	.5010	.3115
		30	.0295	.0790	.0295

[a]Nominal level of .05 for each hypothesis.

[b]Nominal level of .05/*k* for each hypothesis; *k* is the number of tests in any one experiment.

[c]Nominal level for each hypothesis is not known.

Table 6

(Continued)

Design	f	n	Error Rates		
			Alpha	Bonferroni	Hartley
2x5	.10	5	.9975	1.0000	.9985
		15	.9955	1.0000	.9990
		30	.9860	.9985	.9885
	.25	5	.9795	.9985	.9855
		15	.7875	.9175	.8015
		30	.3720	.5645	.3770
	.40	5	.8450	.9395	.8575
		15	.2045	.3765	.2050
		30	.0165	.0435	.0165
5x5	.10	5	.9970	.9995	.9990
		15	.9855	.9960	.9880
		30	.9225	.9805	.9335
	.25	5	.9165	.9800	.9305
		15	.3500	.5260	.3545
		30	.0260	.0700	.0260
	.40	5	.5040	.6985	.5140
		15	.0065	.0225	.0065
		30	.0000	.0000	.0000
2x2x2	.10	5	1.0000	1.0000	1.0000
		15	1.0000	1.0000	1.0000
		30	1.0000	1.0000	1.0000
	.25	5	1.0000	1.0000	1.0000
		15	.9730	1.0000	.9865
		30	.7750	.9755	.8060
	.40	5	.9905	1.0000	.9950
		15	.6060	.9295	.6330
		30	.1605	.4530	.1625

Table 7

Experimentwise Error Rates:

2x4 Designs with Various *f*'s

f_A	f_B	f_{AB}	*n*	Type I Alpha[a]	Type I Bonf.[b]	Type I Hartley[c]	Type II Alpha[a]	Type II Bonf.[b]	Type II Hartley[c]
.25	0	0	5	.0970	.0340	.0375	.6620	.8130	.8120
			15	.0905	.0275	.0380	.2290	.3815	.3815
			30	.0950	.0380	.0500	.0290	.0700	.0695
0	.25	0	5	.0995	.0335	.0370	.7860	.9005	.8980
			15	.0925	.0295	.0340	.3945	.5865	.5850
			30	.1050	.0350	.0500	.0800	.1700	.1680
.25	0	.25	5	.0510	.0170	.0210	.9470	.9835	.9760
			15	.0535	.0220	.0315	.7400	.8755	.8375
			30	.0480	.0350	.0365	.4185	.5895	.5325
.40	0	.25	5	.0515	.0155	.0265	.8910	.9680	.9455
			15	.0500	.0195	.0295	.6775	.8215	.7780
			30	.0560	.0145	.0440	.3910	.5540	.4960
0	.25	.25	5	.0495	.0170	.0185	.9625	.9915	.9850
			15	.0410	.0155	.0200	.8065	.9255	.8857
			30	.0575	.0200	.0390	.4380	.6300	.5625
0	.40	.25	5	.0480	.0155	.0215	.9260	.9780	.9675
			15	.0575	.0195	.0335	.6715	.8065	.7575
			30	.0540	.0190	.0430	.3795	.5595	.4930

[a]Nominal level of .05 for each hypothesis.

[b]Nominal level of .05/*k* for each hypothesis; *k* is the number of tests in any one experiment.

[c]Nominal level for each hypothesis is not known.

Table 8

Experimentwise Error Rates:

2x2x2 Designs with Various $\underline{f}$'s

f_A	f_{AB}	f_{ABC}[a]	n	Type I Alpha[b]	Type I Boni[c]	Type I Hartley[d]	Type II Alpha[b]	Type II Boni[c]	Type II Hartley[d]
.25	0	0	5	.2560	.0355	.0360	.6510	.8780	.8775
			15	.2795	.0415	.0450	.2215	.5030	.5020
			30	.2370	.0380	.0435	.0340	.1270	.1265
.25	.25	0	5	.2290	.0340	.0350	.9315	.9910	.9890
			15	.2525	.0375	.0415	.6175	.8895	.8715
			30	.2465	.0385	.0505	.2415	.5620	.5310
.40	.25	0	5	.2350	.0360	.0375	.8450	.9755	.9710
			15	.2280	.0390	.0500	.5145	.7900	.7750
			30	.2375	.0345	.0490	.2045	.4890	.4590
.25	.25	.25	5	.1870	.0215	.0225	.9860	1.0000	.9995
			15	.1715	.0270	.0310	.8825	.9875	.9805
			30	.2160	.0275	.0340	.6215	.8945	.8600
.40	.40	.25	5	.1920	.0280	.0355	.9590	.9965	.9945
			15	.1780	.0265	.0380	.7535	.9400	.9165
			30	.1975	.0305	.0505	.5130	.7960	.7575

[a]All additional effects set at zero; that is, $f_B = f_C = f_{AC} = f_{BC} = 0$.

[b]Nominal level of .05 for each hypothesis.

[c]Nominal level of .05/k for each hypothesis; k is the number of tests in any one experiment.

[d]Nominal level for each hypothesis is not known.